2017 25th Mediterranean Conference on Control and Automation (MED)
July 3-6, 2017. Valletta, Malta

Multi-gated Perimeter Flow Control of Transport Networks*

Ruzanna Mat Jusoh and Konstantinos Ampountolas

Abstract— **This paper develops a control scheme for the multi-gated perimeter traffic flow control problem of urban road networks. The proposed scheme determines optimally distributed input flows (or feasible entrance link green times) for a number of gates located at the periphery of a protected network area. A macroscopic model is employed to describe the traffic dynamics of the protected network. To describe traffic dynamics outside of the protected area, we augment the basic state-space model with additional state variables to account for the queues at store-and-forward origin links at the periphery. We aim to equalise the relative queues at origin links and maintain the vehicle accumulation in the protected network around a desired point, while the system's throughput is maximised. The perimeter traffic flow control problem is formulated as a convex optimal control problem with constrained control and state variables. For real-time control, the optimal control problem is embedded in a rolling-horizon scheme using the current state of the whole system as the initial state as well as predicted demand flows at entrance links. A meticulous simulation study is carried out for a 2.5 square mile protected network area of San Francisco, CA, including fifteen gates of different geometric characteristics. Results demonstrate the efficiency and equity properties of the proposed approach to better manage excessive queues outside of the protected network area and optimally distribute the input flows.**

I. Introduction

Traffic congestion on urban networks is deemed to inefficient road operations and excessive traffic demand, which calls for drastic solutions. The performance of road infrastructure is usually assessed by microscopic models at the link and/or junction level. In an attempt to assess the performance of urban networks at a macroscopic level, a parsimonious but not accurate model is often used, which primarily shows the relationship between average network flow and vehicle accumulation or density. A macroscopic model of steady-state urban traffic was proposed by [1], [2], further developed by [3], [4], [5], [6] and fitted to experimental data by [7], [8] and others. This model is the so-called Macroscopic or Network Fundamental Diagram (MFD or NFD) of urban road networks; it presumes (under certain regularity conditions) that traffic flows dynamics can be treated macroscopically as a single-region dynamic system with vehicle accumulation n as a single state variable. The main feature of the NFD (with a concave like-shape as in Fig. 1) is that for a critical vehicle accumulation $\hat{n}$ flow capacity is reached (maximum throughput). This property can be utilised to introduce perimeter flow control policies to improve mobility in single-region homogeneous [4], [9], [10] or multi-region heterogeneous networks [11], [12], [13], [14], [15]. A perimeter flow control policy "meters" the input flow to the system and hold vehicles outside a protected network area if necessary, so as to maximise the throughput.

Except of a few works such as [16], [17], [18], studies on the perimeter flow control assume that a single ordered input flow is equally distributed to a number of candidate junctions at the periphery of the network without taking into account the different geometric characteristics of origin links, i.e., length, storage capacity, etc. Such a distribution applied independently to multiple gates of a protected network area would be efficient in case of unconstrained origin link queues for vehicle storage. However, gated queues must be restricted to avoid interference with adjacent street traffic outside of the protected network area and geometric characteristics of the different gates must be taken into account in the optimisation. Thus, limited origin links storage capacity, geometric characteristics and the requirement of equity for drivers using different gates to enter a protected area are the main reasons towards multi-gated perimeter flow control.

We propose an integrated model for the multi-gated perimeter traffic flow control problem in cities. We employ the network or macroscopic fundamental diagram of urban networks to describe the traffic dynamics of the protected network area. To describe traffic dynamics outside of the protected area, we augment the basic state-space model with additional state variables for the queues at store-and-forward origin links at the periphery. This model is then used to formulate a convenient convex control problem with constrained control and state variables for multi-gated perimeter flow control. This scheme determines optimally distributed input flow values to avoid queues and delays at the perimeter of a protected area while system's output is maximised. For the convex optimisation problem, we present results to demonstrate its efficiency and equity properties to better manage excessive queues outside of the protected network area and optimally distribute input flows.

II. Modelling with Entrance Links Dynamics

Consider a protected network area with a number of controlled gates $o \in \mathcal{O} = \{1, 2, \ldots\}$ located at its periphery as shown in Fig. 1. The set $\mathcal{O}$ includes all the origin links whose outflow is essentially entering into the protected network from a number of controlled gates/entrances (e.g. signalised junctions or toll stations). In principle, the origin links at the periphery of the protected network would have different geometric characteristics, i.e., length, number of lanes, capacity, saturation flows.

*The first author is supported by the Malaysian Government and the National Defense University (NDUM) of Malaysia.

Ruzanna Mat Jusoh and Konstantinos Ampountolas are with the School of Engineering, University of Glasgow, Glasgow G12 8QQ, United Kingdom. E-mail: R.Mat-Jusoh.1@research.gla.ac.uk and Konstantinos.Ampountolas@glasgow.ac.uk

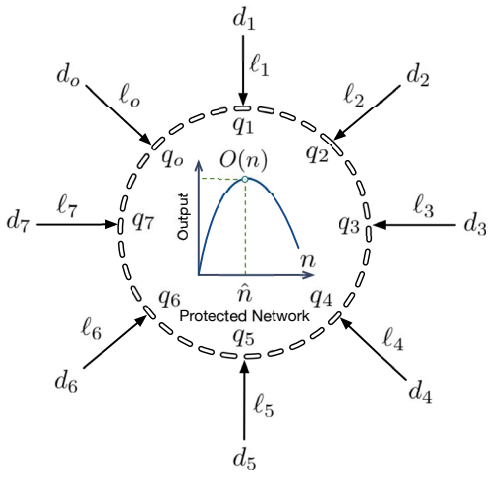


Fig. 1. Protected network with entrance link dynamics.

The protected network traffic can be treated macroscopically as a single-region dynamic system with vehicle accumulation $n(t)$, $t \geq 0$ as a single state variable [4]. Assume there exists a well-defined function $O(n(t))$ (veh/h) that provides the estimated flow (output) at which vehicles complete trips per unit time either because they finish their trip within the network or because they move outside of the network. This function describes steady-state behaviour of single-region homogeneous networks if the input to output dynamics are not instantaneous and any delays are comparable with the average travel time across the network. The output (throughput) function $O(n(t))$ of a network can be easily determined if trip completion rates or Origin-Destination (OD) data are available in real-time (e.g. from vehicles equipped with GPS trackers). Alternatively, the output can be expressed as $O(n(t)) = (l/L)O_c(n(t))$, where L (m) is the average trip length in the network, l (m) is the average link length, and O_c (veh/h) is the total network circulating flow. In general, the circulating flow O_c can be estimated if $n(t)$ is observed in real-time.

Let $q_o(t)$ (veh/h) be the outflow of gate $o \in \mathcal{O}$ at time t. Also, let $q_{\text{out}}(t)$ (veh/h) and $d_n(t)$ (veh/h) be the total outflow and the uncontrolled traffic demand (disturbances) of the protected network at time t, respectively. Note that $d_n(t)$ includes both internal (off-street parking for taxis and private vehicles) and external (from the periphery) non-controlled inflows. The dynamics of the system are governed by the following nonlinear conservation equation

$$\dot{n}(t) = \sum_{o=1}^{|\mathcal{O}|} q_o(t - \tau_o) - q_{\text{out}}(t) + d_n(t), \tag{1}$$

where $q_{\text{out}}(t)$ is in general a nonlinear function of vehicle accumulation $n(t)$ and τ_o is the travel time needed for vehicles to approach the protected network area from origin link $o \in \mathcal{O}$. The time lags τ_o may be translated into an according number of time steps for a discrete-time representation, provided a closed system with inflows q_o and outflow q_{out}. Without loss of generality, we assume that $\tau_o = 0$, $\forall\, o \in \mathcal{O}$, i.e., vehicles released from the controlled gates can immediately get access to the protected network. Moreover, since the system evolves slowly with time t, we may assume that outflow $q_{\text{out}}(t) \propto O_c(n(t))$, and it may thus be given in terms of the output $O(n(t))$. Note that $q_o(t)$, $o \in \mathcal{O}$ are the input variables of the controlled gates/entrances, to be calculated by a multi-gated perimeter flow control strategy.

To describe traffic dynamics outside of the protected area, we augment the basic state-space model (1) with additional state variables for the queues at store-and-forward entrance links at the periphery. Each origin link ℓ_o receives traffic demand d_o and forward it into the protected network, as shown in Fig. 1. The queuing model for the entrance link dynamics is described by the following conservation equation

$$\dot{\ell}_o(t) = d_o(t) - q_o(t), \quad o \in \mathcal{O} = \{1, 2, \ldots\} \tag{2}$$

where $\ell_o(t)$ (veh) and $d_o(t)$ (veh/h) are the vehicle queue and traffic demand in origin link o at time t, respectively.

The integrated model (1)–(2) can be extended to consider a broader class of state and control constraints. For example, inequality state and control constraints may be introduced to preserve congested phenomena within the protected network and to avoid long queues and delays at the perimeter of the network where gating is literally applied. These constraints may be brought to the form

$$\begin{aligned} &0 \leq n(t) \leq n_{\max} \\ &0 \leq \ell_o(t) \leq \ell_{o,\max}, \quad o \in \mathcal{O} = \{1, 2, \ldots\} \\ &q_{o,\min} \leq q_o(t) \leq q_{o,\max}, \quad o \in \mathcal{O} = \{1, 2, \ldots\} \end{aligned} \tag{3}$$

where $n_{\max}$ is the maximum vehicle accumulation of the protected network; $\ell_{o,\max}$ is the maximum permissible capacity of link $o \in \mathcal{O}$; $q_{o,\min}$, $q_{o,\max}$ are the minimum and maximum permissible outflows, respectively; and, $q_{o,\min} > 0$ to avoid long queues and delays at the periphery of the network. Link capacities and maximum vehicle accumulation depend on geometric characteristics of the origin links (length, number of lanes) and the topology of the protected network, respectively. Minimum and maximum permissible outflows can easily be determined given saturation flows, minimum and maximum green times, and cycle times of a nominal traffic signal plan (or corresponding toll ticket) at each controlled gate of the protected network.

The presented model can be viewed as a nonlinear process with input variables $\mathbf{u}^{\mathsf{T}} = \begin{bmatrix} q_1 & q_2 & \cdots & q_{|\mathcal{O}|} \end{bmatrix}$, state variables $\mathbf{x}^{\mathsf{T}} = \begin{bmatrix} n & \ell_1 & \ell_2 & \cdots & \ell_{|\mathcal{O}|} \end{bmatrix}$, and disturbances $\mathbf{d}^{\mathsf{T}} = \begin{bmatrix} d_n & d_1 & d_2 & \cdots & d_{|\mathcal{O}|} \end{bmatrix}$. Then, the continuous-time nonlinear state system (1), (2) with constraints (3) for a protected network with controlled gates $o \in \mathcal{O}$, may be rewritten in compact vector form as

$$\dot{\mathbf{x}}(t) = \mathbf{f}\left[\mathbf{x}(t), \mathbf{u}(t), \mathbf{d}(t), t\right], \, t \geq 0, \quad \mathbf{x}(0) = \mathbf{x}_0 \tag{4}$$

$$\mathbf{0} \leq \mathbf{x}(t) \leq \mathbf{x}_{\max} \tag{5}$$

$$\mathbf{u}_{\min} \leq \mathbf{u}(t) \leq \mathbf{u}_{\max} \tag{6}$$

where $\mathbf{f}$ is a nonlinear vector function reflecting the right-hand side of (1)–(2); $\mathbf{x}_0$ is a known initial state; and $\mathbf{x}_{\max}$, $\mathbf{u}_{\min}$, $\mathbf{u}_{\max}$ are vectors of appropriate dimension reflecting the upper and lower bounds of constraints (3).

Assuming a nonlinear representation of $q_{\text{out}}(t) \triangleq O(n(t))$, the continuous-time nonlinear model (4) may be linearised around some set point $\hat{\mathbf{s}}^{\mathsf{T}} = \begin{bmatrix}\hat{\mathbf{x}} & \hat{\mathbf{u}} & \hat{\mathbf{d}}\end{bmatrix}$, and directly translated into discrete-time, using Euler first-order time discretisation with sample time T, as follows

$$\Delta\mathbf{x}(k+1) = \mathbf{A}\Delta\mathbf{x}(k) + \mathbf{B}\Delta\mathbf{u}(k) + \mathbf{C}\Delta\mathbf{d}(k) \qquad (7)$$

where $k = 0, 1, \ldots, N_o - 1$ is a discrete time index with optimisation horizon N_o; $\Delta(\cdot) \triangleq (\cdot) - \hat{\cdot}$ for all vectors; and $\mathbf{A} = \partial\mathbf{f}/\partial\mathbf{x}|_{\hat{\mathbf{s}}}$, $\mathbf{B} = \partial\mathbf{f}/\partial\mathbf{u}|_{\hat{\mathbf{s}}}$, $\mathbf{C} = \partial\mathbf{f}/\partial\mathbf{d}|_{\hat{\mathbf{s}}}$ are the state, control, and disturbance matrices, respectively. This discrete-time linear model is completely controllable and reachable.

The sample time interval T is literally selected to be a common multiple of cycle lengths of all controlled gates at the periphery of the protected network, while $T \in [3, 5]$ min is usually appropriate for constructing a well-defined outflow function $O(n(t))$, given experimental data. In principle, origin link dynamics (2) are much faster than the dynamics of the protected network (1) (governed by the NFD, which evolves slowly in time). Therefore two different time steps T_n and T_ℓ (where $T_\ell \ll T_n$) can be employed for (1) and (2), respectively, to account storage capacity and dispersion of the flow phenomena within store-and-forward origin links; and thus increase model accuracy.

III. Multi-Gated Perimeter Flow Control

A. Control Objective

A natural control objective is to minimise the total time that vehicles spend in the system including both time waiting at origin links to enter and time traveling in the protected network. However, such a policy may induce unbalanced gating of vehicles at the origin links of the protected network, and, as a consequence, would lead to long queues and overflow phenomena within origin links. Unbalanced gating would also violate the requirement of equity for drivers using different gates to enter a protected network area.

Given these observations, a suitable control objective for a protected network area with origin links queue dynamics aims at: (a) equalising the relative vehicle queues $\ell_o/\ell_{o,\max}$, $o \in \mathcal{O}$ over time, and (b) maintaining the vehicle accumulation in the protected network around a set (desired) point $\hat{n}$ while the system's throughput is maximised. A quadratic criterion that considers this control objective has the form

$$J = \frac{1}{2}\sum_{k=0}^{N_o-1}\Big(\|\Delta\mathbf{x}(k)\|^2_{\mathbf{Q}} + \|\Delta\mathbf{u}(k)\|^2_{\mathbf{R}}\Big) \qquad (8)$$

where $\mathbf{Q}$ and $\mathbf{R}$ are positive semi-definite and positive definite diagonal weighting matrices, respectively. The diagonal elements of Q (see definition of vector $\mathbf{x}$) are responsible for balancing the relative vehicle accumulation of the protected network $n/n_{\max}$ and the relative vehicle queues $\ell_o/\ell_{o,\max}$, $o \in \mathcal{O}$. Given that vehicle storage in the protected network is significantly higher than in the origin links, a meticulous selection of diagonal elements is required. A practicable choice is to set $\mathbf{Q} = \text{diag}(1/w, 1/\ell_{1,\max}, \ldots, 1/\ell_{|\mathcal{O}|,\max})$, where the scale of $w \ll n_{\max}$ is of the order of $\sum_{o=1}^{|\mathcal{O}|}\ell_{o,\max}$ to achieve equity. It becomes quite clear here that equity at origin links and efficiency of the protected network area are partially competitive criteria, hence a perimeter flow control strategy should be flexible enough to accommodate a particular trade-off (i.e. to give priority to the protected network or the outside area, e.g. to manage better excessive queues) to be decided by the responsible network authorities. Finally, the choice of the weighting matrix $\mathbf{R} \triangleq r\mathbf{I}$, $r > 0$ can influence the magnitude of the control actions and thus r should be selected via a trial-and-error process.

B. Rolling-Horizon Control

Rolling-horizon control is a repetitive optimisation scheme, where at each time step an open-loop optimal control problem with finite horizon N_o and predicted demands $\mathbf{d}(k)$ over a prediction horizon N_p is optimised; then only the first control move is applied to the plant and the procedure is carried out again. This rolling-horizon procedure closes the loop that is avoids myopic control actions, while embedding a dynamic open-loop optimisation problem in a closed-loop structure. Predicted demand flows may be calculated by use of historical information or suitable extrapolation methods.

Given the known initial state $\mathbf{x}(0) = \mathbf{x}_0$, a static convex optimisation problem may be formulated over N_o due to the discrete-time nature of the involved process. To see this, assume $N_o = N_p$ and define the vectors:

$$\begin{aligned}
\Delta\mathbf{X} &= \begin{bmatrix}\Delta\mathbf{x}(1)^{\mathsf{T}} & \Delta\mathbf{x}(2)^{\mathsf{T}} & \cdots & \Delta\mathbf{x}(N_o)^{\mathsf{T}}\end{bmatrix}^{\mathsf{T}}\\
\Delta\mathbf{U} &= \begin{bmatrix}\Delta\mathbf{u}(0)^{\mathsf{T}} & \Delta\mathbf{u}(1)^{\mathsf{T}} & \cdots & \Delta\mathbf{u}(N_o-1)^{\mathsf{T}}\end{bmatrix}^{\mathsf{T}}\\
\Delta\mathbf{D} &= \begin{bmatrix}\Delta\mathbf{d}(0)^{\mathsf{T}} & \Delta\mathbf{d}(1)^{\mathsf{T}} & \cdots & \Delta\mathbf{d}(N_p-1)^{\mathsf{T}}\end{bmatrix}^{\mathsf{T}}.
\end{aligned}$$

Assuming now availability of demand flow predictions at the origin links of the protected network over a prediction horizon N_p, i.e. $\Delta\mathbf{d}(k) \neq \mathbf{0}$, $k = 0, 1, \ldots, N_p - 1$, minimisation of the performance criterion (8) subject to (7) leads to the analytical solution:

$$\Delta\mathbf{U} = -\mathbf{H}^{-1}\mathbf{F}\big[\mathbf{x}_0 + \mathbf{G}\Delta\mathbf{D}\big], \qquad (9)$$

where $\mathbf{H} = \mathbf{\Gamma}^{\mathsf{T}}\mathcal{Q}\mathbf{\Gamma} + \mathcal{R}$ is the Hessian of the corresponding quadratic program (QP), $\mathbf{F} = \mathbf{\Gamma}^{\mathsf{T}}\mathcal{Q}\mathbf{\Omega}$, and $\mathbf{G} = \mathbf{\Gamma}^{\mathsf{T}}\mathcal{Z}$. The matrices $\mathbf{\Gamma}$ and $\mathbf{\Omega}$ may be readily specified from the integration of (7) starting from the initial point $\mathbf{x}(0)$, while $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{Z}$ are weighting matrices (in function of $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{C}$) over the optimisation N_o [19]. Given that $\mathbf{R} \succ \mathbf{0}$ in the cost criterion (8) the Hessian $\mathbf{H}$ is positive definite, and thus the QP is convex and has a global optimum. The third term may be regarded as a feedforward term, accounting for future disturbances. For $N_o \to \infty$ and vanishing disturbances, i.e., $\Delta\mathbf{d}(k) = \mathbf{0}$, $\forall\, k = 0, 1, \ldots, N_p - 1$, a Linear-Quadratic-Integral regulator may be derived as in [12].

Using the above formalism, we can express the problem of minimising (8) subject to the equality constraints (7) and inequality constraints (5)–(6) as follows:

$$\begin{aligned}
&\min_{\mathbf{U}}\ \frac{1}{2}\mathbf{U}^{\mathsf{T}}\mathbf{H}\mathbf{U} + \mathbf{U}^{\mathsf{T}}\big[\mathbf{F}\mathbf{x}(0) - \mathbf{H}\hat{\mathbf{U}} + \mathbf{G}\Delta\mathbf{D}\big]\\
&\text{subject to:}\quad \mathbf{L}\mathbf{U} \leq \mathbf{W}
\end{aligned} \qquad (10)$$

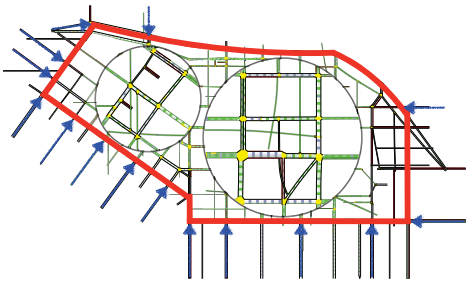


Fig. 2. Protected network and controlled gates of entrance.

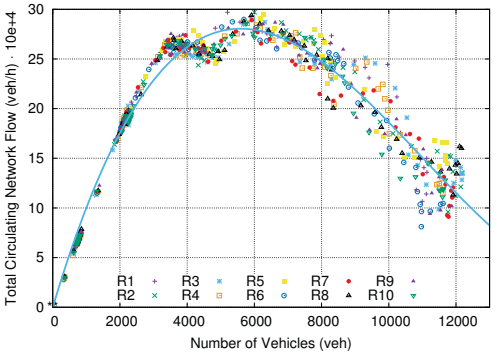


Fig. 3. Network fundamental diagram of Downtown SF.

where $\mathbf{L}$ and $\mathbf{W}$ are matrices reflecting the lower and upper bounds of the state and control constraints (given state integration starting from the point $\mathbf{x}_0$) over the optimisation horizon N_o [19]. Once the open-loop QP problem (10) is solved from the known initial $\mathbf{x}(0)$ and predicted disturbances $\mathbf{d}(k), k = 0, 1, \ldots, N_p - 1$, the rolling horizon scheme applies, at the current time k, only the first control move, formed by the first m components of the optimal vector $\mathbf{U}^*(\mathbf{x}_0)$ in (10). This yields a control law of the form $\mathbf{u}(k) = \mathcal{M}\big[\mathbf{x}(k), \mathbf{d}(\kappa)\big], \kappa = k, k+1, \ldots, k+N_p-1$, where $\mathbf{x}(k) = \mathbf{x}_0$, $k = 0, \ldots, N_o - 1$ is the current state of the system and $\mathcal{M}$ is a linear mapping from the state and disturbance spaces to control. Then the whole procedure is repeated at the next time instant.

IV. Application and Results

A. Network Description and Controller Design

Fig. 3 depicts the shape of O_c in function of $n(t)$ for the 2.5 square mile area of Downtown San Francisco (SF), CA, including 110 junctions and 440 links (see Fig. 2). Fig. 3 confirms the existence of an NFD like-shape for the study area, which shape depends on $n(t)$. It can be seen that as the vehicle accumulation is increased from zero, the network flow increases to a maximum (flow capacity) and then turns down and decreases sharply to a low value possibly zero (in case of gridlock). Capacity (around $30{\cdot}10^4$ veh/h) is observed at $n(t)$ of about 6000 veh. The shape of the NFD was reproduced under different demand scenarios with Dynamic Traffic Assignment activated to capture adaptive drivers in microsimulation via AIMSUN [12]. The shape of O_c in Fig. 3 can be approximated by the 3rd order polynomial:

$$O_c(n) = 4.128 \times 10^{-7} n^3 - 0.0136 n^2 + 113.264 n, \quad (11)$$

where $n \in [0, 13000]$ veh. To determine O from O_c an average trip length $L = 1.75$ km and average link length $l = 0.25$ km were considered. The value of L is consistent with the average trip length across the test area of SF.

The desired vehicle accumulation for (7) is selected $\hat{n} = 4000$ veh, while $\hat{\ell}_o = 0$, $\forall o \in \mathcal{O}$. Table I provides the different geometric characteristics of the fifteen ($|\mathcal{O}| = 15$) controlled gates shown in Fig. 2 (illustrated with blue arrows), thus $\mathbf{x} \in \mathbb{R}^{16}$ and $\mathbf{u} \in \mathbb{R}^{15}$. The third column provides the storage capacity of each controlled link that is the vector $\mathbf{x}_{\max}^{\mathsf{T}} = \begin{bmatrix} n_{\max} & \ell_{1,\max} & \cdots & \ell_{|\mathcal{O}|,\max} \end{bmatrix} \in \mathbb{R}^{16}$. The last

TABLE I

Geometric Characteristics of Controlled Gates

Gate	Length	Capacity	# lanes	Saturation Flow	Cycle Length	Min Green	Nominal Green	Max Green	Min Flow	Nominal Flow	Max Flow
#	(m)	(veh)		(veh/h)	(s)	(s)	(s)	(s)	(veh/h)	(veh/h)	(veh/h)
1	235	128	3	5400	60	15	33	39	1350	2970	3510
2	299	109	2	3600	60	12	30	42	720	1800	2520
3	299	163	3	5400	60	15	27	39	1350	2430	3510
4	271	98	2	3600	60	12	35	42	720	2100	2520
5	261	95	2	3600	60	12	24	42	720	1440	2520
6	299	109	2	3600	60	12	30	42	720	1800	2520
7	298	109	2	3600	60	12	36	39	720	2160	2340
8	298	109	2	3600	60	12	37	42	720	2220	2520
9	296	269	5	10000	60	17	27	31	2833	4500	5167
10	296	269	5	8800	60	16	27	38	2347	3960	5573
11	299	109	2	3600	60	12	25	42	720	1500	2520
12	190	103	3	5400	90	13	42	43	780	2520	2580
13	81	44	3	5400	90	12	18	41	720	1080	2460
14	81	44	3	5400	90	12	20	41	720	1200	2460
15	341	186	3	5400	90	12	48	59	720	2880	3540

three columns of the table provide the vectors $\mathbf{u}_{\min} = \mathbf{q}_{\min}$, $\hat{\mathbf{u}} = \hat{\mathbf{q}}$, and $\mathbf{u}_{\max} = \mathbf{q}_{\max}$, respectively. These values are calculated from the field applied signal plans presented in columns 5 (S_o: saturation flow), 6 (C: cycle length) , 7 ($g_{o,\min}$: min green time), 8 ($\hat{g}_o$: nominal green time), and 9 ($g_{o,\max}$: max green time). In this way, any input flows ordered by the multi-gated perimeter flow control strategy are feasible traffic signal plans. Traffic signals at controlled gates are all multiphase operating on a common cycle length of 90 s for the west boundary (gates $o = 12, \ldots, 15$) and 60 s for the rest (gates $o = 1, \ldots, 11$).

For the solution of (9) it suffices to specify the state matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, and weighting matrices $\mathbf{Q}$ and $\mathbf{R}$. All state matrices are constructed for the studied network on the basis of the selected $\hat{\mathbf{x}}^{\mathsf{T}} = [\hat{n} \quad \mathbf{0}] \in \mathbb{R}^{16}$, $\hat{\mathbf{u}} = \hat{\mathbf{q}} \in \mathbb{R}^{15}$ and $\hat{\mathbf{d}} = \mathbf{0}$, and sampling time $T = 180$ s. More precisely, $\mathbf{A} = \mathrm{diag}(1 - 4.94 \times T, 1, \ldots, 1) \in \mathbb{R}^{16\times16}$, $\mathbf{B} = T\begin{bmatrix}\mathbf{1}_{1\times15} & -\mathbf{I}_{15\times15}\end{bmatrix}^{\mathsf{T}} \in \mathbb{R}^{16\times15}$, and $\mathbf{C} = \mathrm{diag}(T, \ldots, T) \in \mathbb{R}^{16\times16}$. The matrix $\mathbf{Q} = \mathrm{diag}(1/w, 1/\ell_{1,\max}, \ldots, 1/\ell_{|\mathcal{O}|,\max})$ is selected, where $w = 2000$ was found appropriate to achieve equity. The diagonal elements of $\mathbf{R}$ were set equal to $r = 0.00001$. The disturbance vector $\mathbf{d}$ consists of the demands d_o, $o = 1, \ldots, 15$, at every origin of the protected network and disturbance d_n of the NFD. Trapezoidal demands have been used for $d_o(k)$, $o = 1, \ldots, 15$, $k = 0, \ldots, N_p - 1$ over a predicted horizon of $N_o = N_p = 40$. To capture the uncertainty of the NFD, particularly when the network operating in the congested regime (notice the noise for $n > 6000$ veh), d_n is selected to vary gradually with respect to $n(k)$ in the range $[-3000, 3000]$ veh/h for $n > 6000$ veh.

B. Control Results and Comparisons

Several scenarios were created by assuming more or less high initial queues $\ell_o(0)$ in the fifteen origin links of the protected network while the protected network area operating in the congested regime, i.e. its state $n(0) > 6000$ veh. The optimisation horizon for each scenario is 2 h (40 cycles). Fig. 4 shows some obtained trajectories for a heavy scenario with $\ell_o(0) = 0.7\ell_{o,\max}, \forall\, o = 1, \ldots, 15$ and initial states in the congested regime of the NFD $n(0) = 7000$ veh and $n(0) = 12000$ veh. Tests were conducted with and without external demand flows at origin links, denoted with "s+d" and "s-d", respectively. The main observations are as follows:

- The proposed multi-gated control strategy manages to stabilise the vehicle accumulation of the protected network around its desired point $\hat{n} = 4000$ veh for all initial points and cases with and without disturbances (see Fig. 4(a)).
- The proposed multi-gated control strategy manages to dissolve the initial origin link queues in a balanced way (see Figs. 4(j)–4(o)) and thus, the desired control objective of queue balancing and equity for drivers using different gates to enter the protected network area is achieved.
- The proposed strategy manages to stabilise all input flows to their desired values $\hat{\mathbf{q}}$ in the steady state (nominal signal plan in Table I), i.e., where $n = \hat{n} = 4000$ (Figs. 4(d)–4(i), notice the different reference points $\hat{q}_o$).
- The input flows ordered by the proposed strategy have different trajectories and characteristics (see control trajectories in Figs. 4(d)–4(i)). This confirms that an equal distribution of ordered flows to corresponding junctions is not optimal, as largely assumed in previous studies. As can be seen, the proposed strategy determines optimally distributed input flows by taking into account the individual geometric characteristics of the origin links as well as minimum and maximum operational constraints.
- It is evident that excessive demand and high initial queues at origin links, coupled with the applied control, causes congestion shortly after the beginning of the time horizon. At the same time the protected network is operating in the congested regime ($n(0) = 7000$ veh or $n(0) = 12000$ veh). As can be seen, the multi-gated control first restricts the high initial queues at origin links to flow into the oversaturated protected network area, and then, it gradually increases the input flows to manage the developed long queues therein. Note that for some gates (7, 8 and 9) bound constraints are active for a certain time period.

Figs. 4(b) and 4(c) depict the control and state trajectories for the perimeter control problem without queue dynamics [12], i.e. for the single-input single output control problem with only (1). As can be seen, this strategy manages to stabilise $n(k)$ of the protected network around its desired point $\hat{n} = 4000$ veh starting from different initial points (including the extreme case of partial gridlock). The strategy restricts flow to enter the protected network area whenever $n > 4000$, while increases the input flows for $n < 4000$. It is evident that the single-region control strategy without queue dynamics outside of the protected area needs less time and effort to stabilise the system at $k = 15$, compared to the proposed multi-gated perimeter flow control, which stabilises all queues and protected network's accumulation at $k = 20$.

Figs 4(p)–4(r) demonstrate the equity properties of the proposed rolling-horizon approach to better manage excessive queues outside of the protected area and optimally distribute the input flows with respect to geometric characteristics. These figures depict results obtained for two different demand scenarios, namely medium and high, and four initial states $n(0)$ of the NFD. As can be seen in Fig. 4(p), Total Time Spent (TTS) within the protected area increases with vehicle accumulation for both scenarios. Figs 4(q)-4(r) demonstrate the equity properties of multi-gated perimeter flow control. Gates with similar characteristics experience similar TTS for two different demand scenarios and four initial states in the protected area. We can distinguish three groups of gates with similar TTS, Group A including gates 2, 4–8, 11–14; Group B including gates 1, 3, 15; and Group C including gates 9, 10. Contrasting gates in Groups A, B, and C with the geometric characteristics in Table I, further supports the equity properties of the proposed control. The control efficiency of the proposed control while explicitly considering queue dynamics and constraints underlines the clear superiority of appropriate multi-gated flow control.

V. CONCLUSIONS

An integrated model for multi-gated perimeter flow control is presented. Compared to previous works, the proposed scheme determines optimally distributed input flows for a number of gates with heterogeneous characteristics located at the periphery of a protected network. Simulation results for a protected area of downtown San Francisco with fifteen gates of different geometric characteristics were presented. Results demonstrated the efficiency and equity properties of the proposed approach to better manage excessive queues outside of the protected network area and optimally distribute the input flows compared to single-region perimeter flow control without queue dynamics. Similar policies can also be utilised for dynamic road pricing.

REFERENCES

[1] J. W. Godfrey, "The mechanism of a road network," *Traffic Engineering and Control*, vol. 11, no. 7, pp. 323–327, 1969.

[2] R. Herman and I. Prigogine, "A two-fluid approach to town traffic," *Science*, vol. 204, pp. 148–151, 1979.

[3] S. Ardekani and R. Herman, "Urban network-wide traffic variables and their relations," *Transport Science*, vol. 21, no. 1, pp. 1–16, 1987.

[4] C. F. Daganzo, "Urban gridlock: Macroscopic modeling and mitigation approaches," *Transport Res Part B*, vol. 41, no. 1, pp. 49–62, 2007.

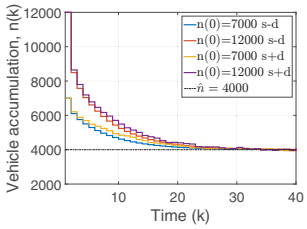
[5] C. F. Daganzo and N. Geroliminis, "An analytical approximation for the macroscopic fundamental diagram of urban traffic," *Transportation Research Part B*, vol. 42, no. 9, pp. 771–781, 2008.

[6] N. Farhi, M. Goursat, and J.-P. Quadrat, "The traffic phases of road networks," *Transport Res Part C*, vol. 19, no. 1, pp. 85–102, 2011.

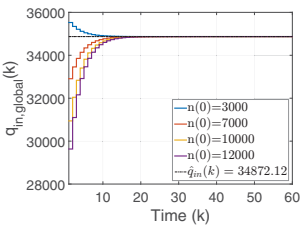
[7] N. Geroliminis and C. F. Daganzo, "Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings," *Transportation Research Part B*, vol. 42, no. 9, pp. 759–770, 2008.

[8] K. Ampountolas and A. Kouvelas, "Real-time estimation of critical vehicle accumulation for maximum network throughput," in *2015 American Control Conference*, 2015, pp. 2057–2062.

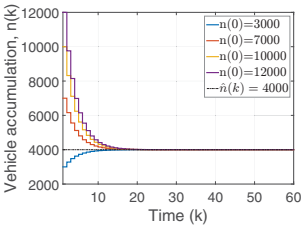
[9] M. Keyvan-Ekbatani, et al., "Exploiting the fundamental diagram of urban networks for feedback-based gating," *Transportation Research Part B*, vol. 46, no. 10, pp. 1393–1403, 2012.



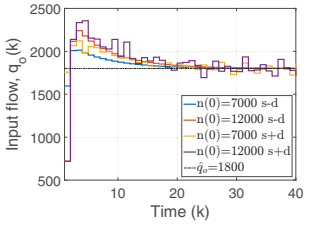
(a) State trajectories (multi-gated control)



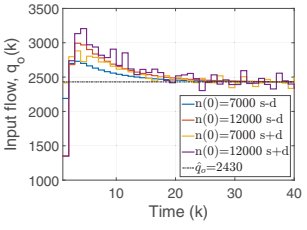
(b) Global input flow



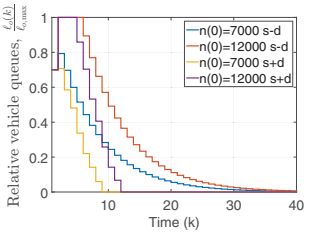
(c) State trajectories (global input)



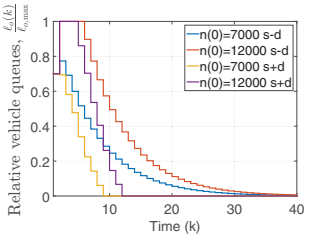
(d) Input flow: Gate 2



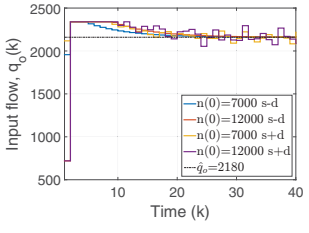
(e) Input flow: Gate 3



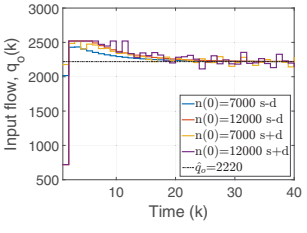
(j) State trajectory: Gate 2



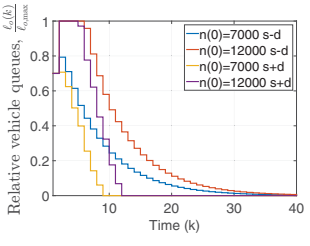
(k) State trajectory: Gate 3



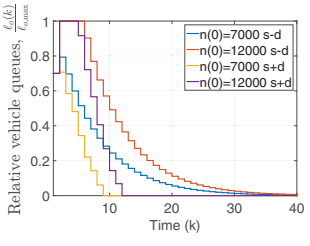
(f) Input flow: Gate 7



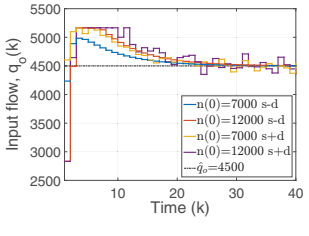
(g) Input flow: Gate 8



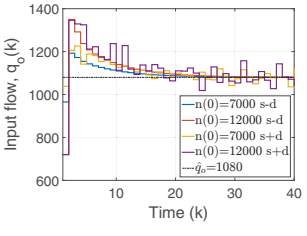
(l) State trajectory: Gate 7



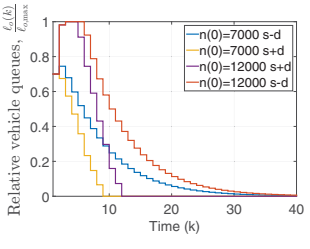
(m) State trajectory: Gate 8



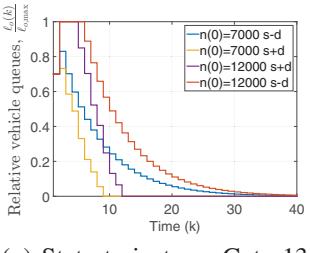
(h) Input flow: Gate 9



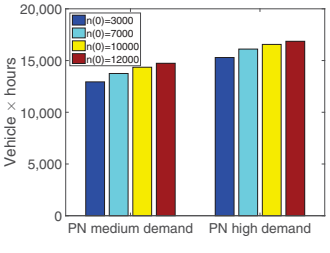
(i) Input flow: Gate 13



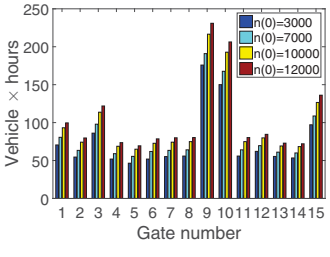
(n) State trajectory: Gate 9



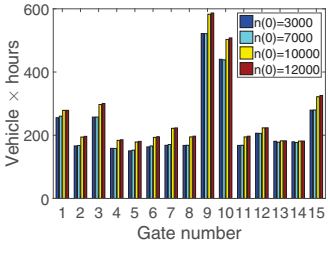
(o) State trajectory: Gate 13



(p) TTS within the protected network area



(q) TTS in gated links (medium demand)



(r) TTS in gated links (high demand)

Fig. 4. (a) State trajectories of the protected network for different initial points with and without disturbance; (b, c) Control and state trajectories without link dynamics; (d–o) Control and state trajectories of six selected gates for different initial points with and without disturbance; (p–r) Total Time Spent (TTS) within the protected network area and in the gated links (outside the protected area).

[10] J. Haddad and A. Shraiber, "Robust perimeter control design for an urban region," *Transport Res Part B*, vol. 68, pp. 315–332, 2014.

[11] N. Geroliminis, et al., "Optimal perimeter control for two urban regions with macroscopic fundamental diagrams," *IEEE Transactions on Intelligent Transport Systems*, vol. 14, no. 1, pp. 348–359, 2013.

[12] K. Aboudolas and N. Geroliminis, "Perimeter and boundary flow control in multi-reservoir heterogeneous networks," *Transportation Research Part B*, vol. 55, pp. 265–281, 2013.

[13] A. Kouvelas, M. Saeedmanesh, and N. Geroliminis, "Enhancing model-based feedback perimeter control with data-driven online adaptive optimization," *Transport Res Part B*, vol. 96, pp. 26–45, 2017.

[14] J. Haddad and B. Mirkin, "Coordinated distributed adaptive perimeter control for large-scale urban road networks," *Transportation Research Part C*, vol. 77, pp. 495–515, 2017.

[15] J. Haddad, "Optimal perimeter control synthesis for two urban regions with aggregate boundary queue dynamics," *Transportation Research Part B*, vol. 96, pp. 1–25, 2017.

[16] A. Csikós, T. Tettamanti, and I. Varga, "Nonlinear gating control for urban road traffic network using the network fundamental diagram," *Journal of Advanced Transportation*, vol. 49, no. 5, pp. 597–615, 2015.

[17] B. Kulcsár, K. Ampountolas, and A. Dabiri, "Single-region robust perimeter traffic flow control," in *2015 European Control Conference*, 2015, pp. 2628–2633.

[18] R. Mat Jusoh and K. Ampountolas, "Distributed perimeter flow control of transport networks," in *49th Annual UTSG Conf*, Dublin, Ireland, 2017.

[19] G. Goodwin, et al., *Constrained Control and Estimation: An Optimisation Approach*. Springer-Verlag, London, UK, 2005.